Detailed, Nationally Consistent Mapping and Spatial Analysis of Section 303(d) and 305(b) Waterbodies using EPA Reach File and National Hydrography Dataset

William S. Cooter, Peter I. Ilieve, Kimberly L. Sparks, Anne Marie Miller,
Christoph Spoerri, and J. Michael McCarthy
Research Triangle Institute

and

Tod Dabolt
Monitoring Branch
U.S. Environmental Protection Agency, Office of Wetlands, Oceans, and Watersheds

ABSTRACT

This paper summarizes a major EPA initiative to locate and map impaired waters across the country using innovative GIS techniques. Results for waters listed by the states in 1998 under Section 303(d) of the Clean Water Act are also presented. Over 18,000 listed waterbodies were located and georeferenced to the EPA Reach File (RF3) hydrography database, which contains locational data for over 3.2 million traces representing surface water features in the continental United States (Dewald et al., 1996). Prior to this initiative, many of the impaired waters listed by the 50 states had never been mapped and could not be displayed on demand or analyzed spatially. Knowing the exact locations and attributes of these listed waters is crucial to developing Total Maximum Daily Loads (TMDLs) and to assessing water quality improvement after TMDLs are implemented.

Using any information each state could provide (hard-copy maps, GIS coverages, Section 305(b) water quality assessment databases, etc.), special GIS coverages were created for the 303(d) listed waters. To achieve national consistency and efficiency, the approach for these coverages involved the creation of an "event" database, which functions as a GIS coverage when used with software such as ESRI's ArcView. A customized program, the Reach Indexing Tool (RIT), was developed as an extension to ESRI's ArcView desktop mapping software. The RIT was made available to the states through training courses and via the Web to enable states to do future reach indexing themselves.

The products of EPA's initiative include detailed maps and GIS coverages, event tables that provide locational data linked to RF3, and conventional relational database files containing information on Section 303(d) pollutants and source categories. Examples of these products, including Web-based materials available to the public, are provided in this paper.

The above technical approaches are now being applied to other major EPA data systems and upgraded to the National Hydrography Dataset (NHD). The newly released NHD replaces RF3 and offers many improved features and capabilities geared to GIS applications. EPA is working with states to georeference their 303(d) lists to NHD; the new NHD version of the RIT is

described in this paper. Also described are recent efforts to map Section 305(b) assessment data, which allows the spatial analysis of healthy as well as impaired waters and state designated uses and criteria, which are crucial to state monitoring, assessment, and management initiatives.

The ultimate goal is to create data systems at the state level that can be easily compiled into uniform national data systems. To this end, EPA also funded development of the Assessment Database, a user-friendly relational database system that states are using to store and analyze their water quality assessment data. The linkage between the GIS coverages and the Assessment Database provides powerful querying and mapping capabilities, as illustrated in this paper. Ultimately, EPA plans to make available these and other types of water quality information through interactive Web-based applications that will allow citizens to select data layers, run database queries, and generate on-screen maps of the results at the high level of spatial resolution offered by the NHD.

INTRODUCTION

States are required to report water quality assessment information and water quality impairments to the U.S. Environmental Protection Agency (EPA) under Sections 305(b) and 303(d) of the Clean Water Act (Cooter et al., 1996, 1998). To meet this requirement, states submit reports in a variety of formats, including databases, spreadsheets, and word processing documents). EPA must then compile these data in a nationally consistent manner. State 305(b) and 303(d) data are rarely accompanied by detailed locational information in the form of maps or GIS coverages. Without a visualization, it is very difficult to observe trends in the national data or identify water quality “hot spots”. Detailed locational information for 303(d) waters is particularly important because it is required for TMDL development.

In 1996, EPA’s Office of Wetlands, Oceans, and Watersheads (OWOW) established an initiative to locate and map impaired waters across the country. The goal was to map these waters in a nationally consistent manner, but also in a format that could easily be used and maintained by the states. To this end, OWOW developed the Reach Indexing Tool (RIT), an interactive ArcView geographic information system (GIS) application that simplifies the process of georeferencing (locating and delineating) state waterbodies to a national hydrography coverage. The RIT has been used to locate and georeference over 18,000 303(d) listed waters to the EPA Reach File (RF3).

Clean Water Act Section 305(b) and Section 303(d) Reporting

Section 305(b) of the Clean Water Act requires states to report on their assessed waters every 2 years. Typically, the state’s surface waters are grouped into “waterbodies” for the purposes of water quality assessment. States report on designated uses, use support, and causes and sources of any impairment for assessed waters for 305(b) reporting. Figure 1 shows a few of New Mexico’s waterbodies.

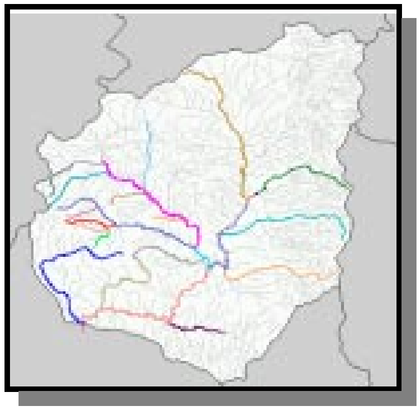

Figure 1. New Mexico 305(b) waterbodies in a USGS 8-digit watershed georeferenced to RF3.

To fulfill the reporting requirements, states submit reports in many different formats, including databases, spreadsheets and word processing documents. These data are compiled and reformatted for inclusion in national databases such as EPA's National Assessment Database (305(b)) or the Total Maximum Daily Load (TMDL) Tracking System (303(d)). To ease the burden of this process for both EPA and the states, OWOW has developed the 305(b) Assessment Database (ADB) (Figure 2). The ADB is an Access application designed to help state users enter and maintain data about their water quality assessments in a nationally consistent manner that can be used with a GIS.

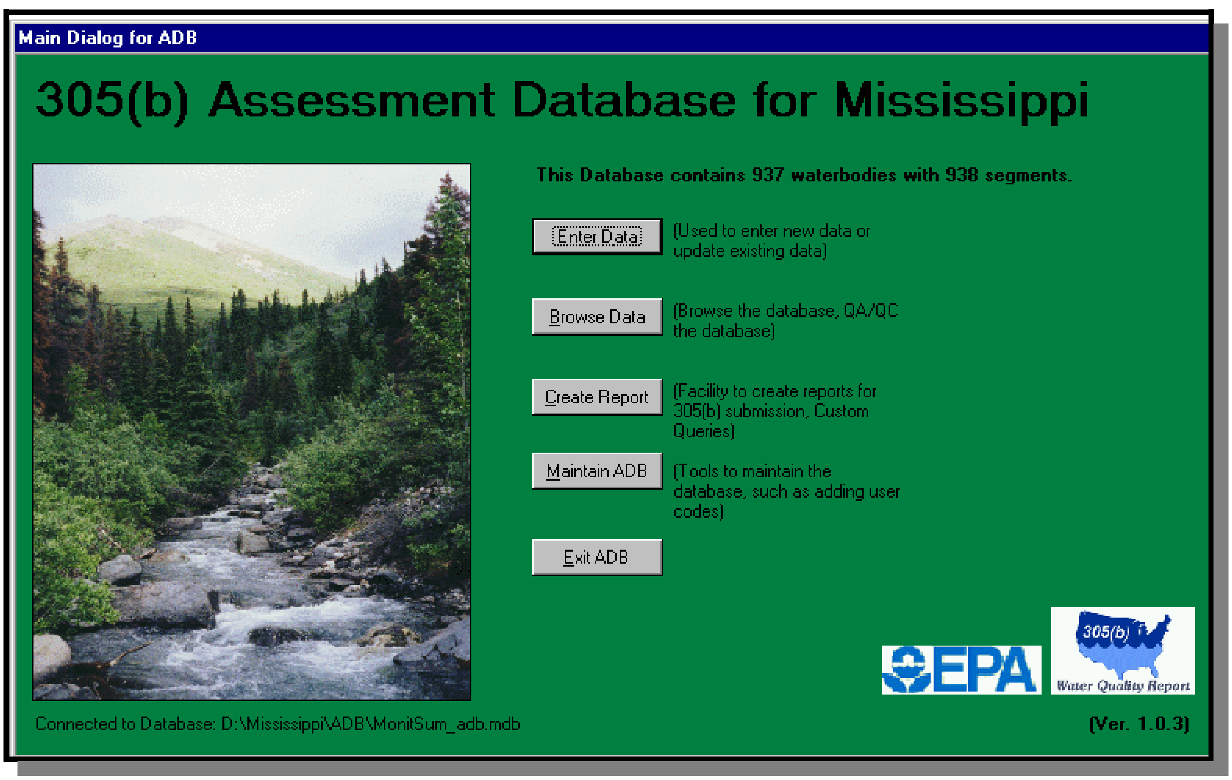

Figure 2. Main Dialog for EPA's 305(b) Assessment Database.

Section 303(d) of the Clean Water Act requires states to identify and establish a priority ranking for waters where existing pollution controls are not stringent enough to attain water quality standards and to establish TMDLs for those waters. Ideally, the 303(d) list is the impaired subset of the waters reported for 305(b).

RF3 and National Hydrography Dataset (NHD)

EPA Reach File (version 3.0), known as RF3, is a national hydrologic database that interconnects and uniquely identifies the 3.2 million stream segments or “reaches” that comprise the country’s surface water drainage system (Dewald et al., 1996). RF3 is designed to serve three simultaneous functions for surface waters:

1. Provide a standard unique identifier (reach code) for each surface water feature (with each feature called a reach)
2. Contain a tabular routing (navigation) network of these features
3. Include a digital map representation of these features.

The reach codes provide a common nomenclature for federal and state reporting of surface water conditions as required under the Clean Water Act. The hydrologic transport network defined in RF3 enables the modeling of waterborne pollution associated with both point and nonpoint sources. The digital map representation of the reaches, combined with the unique reach codes, provides the foundation for detailed, nationally consistent mapping of surface water entities (Figure 3).

RF3 was first compiled in 1992. It is a linear network, containing features from the 1:100,000 USGS digital line graph data. RF3 has been used in many mapping and modeling applications, including a recent initiative by EPA OWOW to map the 1998 303(d) listed waters reported by the states.

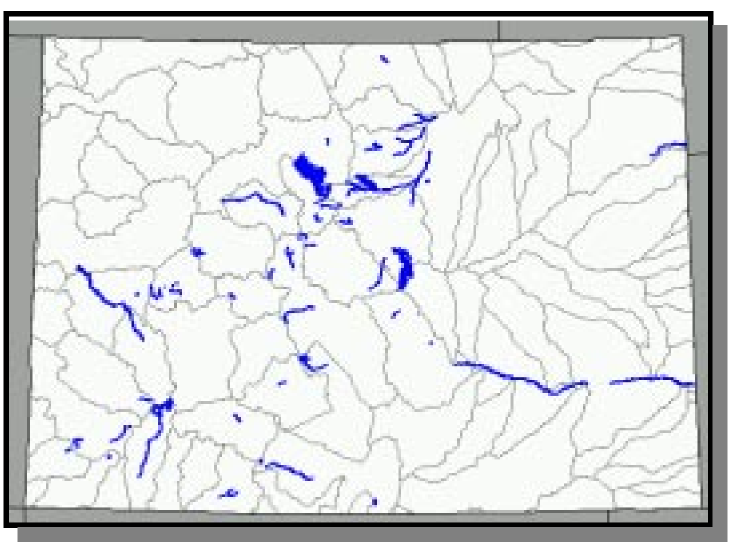

Figure 3. Colorado 1998 303(d) Listed Waters georeferenced to RF3.

The National Hydrography Dataset (NHD) is the next generation reach file. It was created by combining the EPA reach data with updated U.S. Geological Survey (USGS) Digital Line Graph (DLG) hydrography. The result is a hydrography dataset with the functionality of RF3, but the spatial accuracy of DLG. In addition to the hydrography line work, NHD also includes point features that represent entities like gaging stations, wells, springs and seeps, and polygon features to represent lakes, reservoirs, and other two-dimensional features. The route feature in NHD contains transport reaches, which facilitate modeling through the 2-D features (Figure 4).

Figure 4. Artificial paths facilitate modeling through the areal feature layer in NHD.

The benefit of using a reach file like RF3 or NHD for surface water indexing is that the surface water data become anchored on reaches with unique identifiers. This provides nationally consistent locational information, as well as data compatibility among agencies and programs. Other benefits of using NHD are that the maintenance of a national coverage will allow users to benefit from the enhancements of others, and the use of NHD will ensure institutional memory for work performed using the national standard. Using a national standard also ensures a readily accessible replacement for coverages that are inadvertently damaged during processing.

Purpose of this Paper

The purpose of this paper is to discuss the methodology and tools that EPA OWOW has developed and used to encourage states to report and maintain their data in a nationally consistent format. Many considerations must be given to the format of the data to ensure that they can be entered into a database accurately, linked to a GIS, and mapped in a meaningful way. This paper will also highlight some of the data products generated from a recent EPA initiative to georeference 1998 303(d) listed waters to RF3.

METHODOLOGY

Using an Attribute Database with GIS—the Spatial Indeterminacy Problem

One common problem with surface waterbody databases is *spatial indeterminacy*. Spatial indeterminacy occurs when entries in a database apply to only portions of a complete waterbody.

The display of attribute data in a GIS is limited to the resolution of the entities delineated in the associated coverage. For example, if users wanted to display the data in Table 1, they would be limited to the entities in Figure 5. You can identify where the 50-mile WB-A waterbody is located, but the 4-mile portion that partially supports swimming cannot be identified.

Table 1. Sample database illustrating waterbodies with designated uses, which can be fully, partially, or not supporting, depending on water quality

Waterbody ID	Waterbody Length (miles)	Designated Use	Fully Supported (miles)	Partially Supported (miles)	Not Supported (miles)
WB-A	50 mi	Swimming	30 mi	4 mi	10 mi
WB-A	50 mi	Fishing	10 mi	30 mi	4 mi
WB-B	39 mi	Swimming	10 mi	10 mi	19 mi
WB-B	39 mi	Fishing	25 mi	10 mi	0 mi
WB-B	39 mi	Drinking water	0 mi	20 mi	19 mi
WB-C	25 mi	Fishing	20 mi	0 mi	6 mi

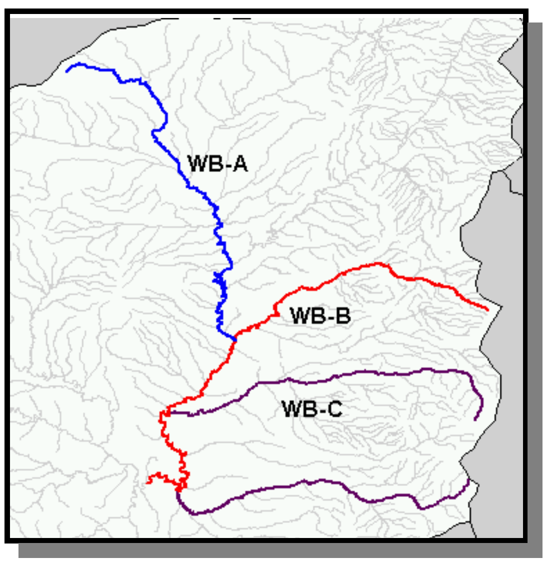

Figure 5. Waterbodies from Table 1 do not have detailed-enough resolution to display the use support information. This limit in resolution is spatial indeterminacy.

This problem is eliminated in EPA's 305(b) Assessment Database through the creation of waterbody segments. Waterbody segments are portions of waterbodies that are treated as homogeneous units, so any designated use support and cause and source data must apply to the waterbody segment in its entirety. Unique identifiers for waterbody segments can be created by concatenating a unique code on the end of the ID. For example, waterbody WB-A in Table 1 could be divided into the waterbody segments in Table 2. These data could be displayed accurately in a GIS, as long as the waterbody segments were delineated in the hydrography coverage.

Table 2. Possible waterbody segments that would allow for accurate mapping of the data pertaining to waterbody WB-A.

Waterbody Segment ID	Waterbody Segment Length (miles)	Swimming Designated Use (miles)	Fishing Designated Use (miles)
WB-A_01	30 mi	Fully supporting	Partially supporting
WB-A_02	10 mi	Not supporting	Fully supporting
WB-A_03	4 mi	Partially supporting	Not supporting

Using a Hydrography Dataset with an Attribute Database

Because NHD has unique identifiers for each reach in the coverage, these numbers can be stored in a database as locational information, much like latitude and longitude are stored. A table containing the list of reach codes that make up a waterbody (or waterbody segment) can be used to display the corresponding waterbody information.

Figure 6 illustrates that if you have a table containing unique identifiers for surface water entities with attribute information (like use support) and you have a table that contains these waterbody identifiers and also the reach codes (RF3RCHID), you can use the reach code information to display the location of the waterbody. Linking waterbody attributes directly to the arc attribute table of a coverage would work moderately well if the surface water entities were composed of entire reach segments. However, surface waters are often described using features that are not part of the hydrography coverage (like bridges and county lines). It would not be desirable to change the hydrography coverage to conform to waterbody descriptions, because then it would no longer be a nationally consistent frame work. Fortunately, ESRI offers a more robust alternative through their Dynamic Segmentation model.

Dynamic segmentation is based on the idea that attributes of a feature can be displayed by simply specifying the start and end points of the feature along an arc, instead of joining the attributes directly to the arc attribute table. The start and end points of these dynamic segments are stored in “To measure” and “From measure” fields in the table containing the reach codes, called an event table. Table 3 illustrates the key fields typically included in a linear event table. This kind of table can be added to an ArcView view as an event theme. Once added to the View, the event theme can be operated on as if it were a coverage.

Waterbody Attribute Information

WBIDSEGID	SEGNAME	USENAME	USECODE	SUPDESC
NC_16-41-1-10_00	Mud Creek	Overall Use Support	1	Fully
NC_16-41-1-11_00	Sandy Creek	Overall Use Support	1	Fully
NC_16-41-1-15-1-(.5)_00	Bolin Creek (Hogan	Overall Use Support	1	Fully
NC_16-41-1-15-1-1-1_00	Buckhorn Branch	Overall Use Support	1	Fully
NC_16-41-1-12-(1)_00	Third Fork Creek	Overall Use Support	1	Not supporting
NC_16-41-1-12-(2)_00	Third Fork Creek	Overall Use Support	1	Not supporting
NC_16-41-1-(14)_00	New Hope Creek	Overall Use Support	1	Partial
NC_16-41-1-15-(3)_00	Little Creek	Overall Use Support	1	Partial
NC_16-41-1-15-(5)_00	Little Creek	Overall Use Support	1	Partial
NC_16-41-1-15-1-(4)_00	Bolin Creek (Hogan	Overall Use Support	1	Partial

LENGTH_M	ENTITY_ID	UPDATE1	RF3RCHID
135.18	NC_16-41-1-15-(.5)_00	081195	30300021727 0.00
119.99	NC_16-41-1-15-(.5)_00	081195	30300021822 0.00
176.48	NC_16-41-1-15-(.5)_00	081195	30300021823 0.00
1106.87	NC_16-41-1-15-(3)_00	081195	30300021446 0.00
6019.53	NC_16-41-1-15-1-(.5)_00	081195	30300021446 6.08
2292.26	NC_16-41-1-15-1-(.5)_00	081195	30300021452 0.00
491.48	NC_16-41-1-15-1-(.5)_00	081195	30300021452 1.42
397.57	NC_16-41-1-15-1-(.5)_00	081195	30300021452 1.73
750.68	NC_16-41-1-15-1-(.5)_00	081195	30300021452 1.97
732.72	NC_16-41-1-15-1-(.5)_00	081195	30300021452 2.44
655.02	NC_16-41-1-15-1-(.5)_00	081195	30300021452 2.45
1300.55	NC_16-41-1-15-1-(.5)_00	081195	30300021452 2.84
338.34	NC_16-41-1-15-1-(.5)_00	081195	30300021452 3.65
357.75	NC_16-41-1-15-1-(.5)_00	081195	30300021452 3.66
1460.54	NC_16-41-1-15-1-(.5)_00	081195	30300021452 3.83
931.08	NC_16-41-1-15-1-(4)_00	081195	30300021446 5.50
654.83	NC_16-41-1-15-1-(4)_00	081195	30300021446 6.08
3139.03	NC_16-41-1-15-1-1_00	081195	30300021455 0.00
8179.74	NC_16-41-1-16-(1)_00	081195	3030002 63 2.17
1764.09	NC_16-41-1-16-(1)_00	081195	3030002 63 0.00

Event Table Georeferenced to RF3

Waterbody Map (RF3)

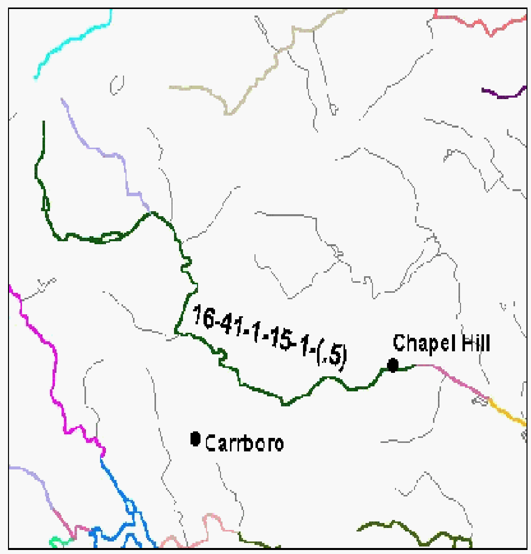

Figure 6. The relationship between waterbody attribute information, an event table and RF3.

Table 3. Fields typically included in an event table

From Measure	To Measure	Reach Code	Entity ID	Offset
From and To measures define the location along the RF3 or NHD reach the event segment should draw.		Unique identifier from RF3 or NHD.	Unique identifier for the entities being mapped (like waterbody segment ID).	Allows multiple event reaches to draw on the same RF3 or NHD reach and display offset from one another.

The use of event tables allows users to index multiple entities to the same NHD reach or set of reaches. It also allows the user to index an entity to a portion of a reach (using the To and From measure fields), without having to modify the spatial extent of the reach itself. This allows for

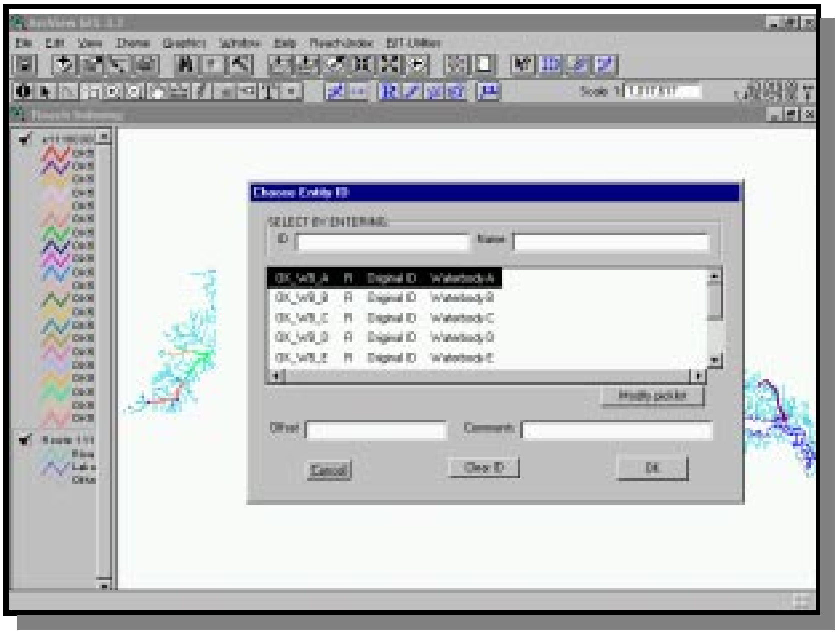

Figure 7. The Reach Indexing Tool adds two new menus and several new buttons and tools to the ArcView interface. After the user has selected a waterbody for delineation, he/she can select an identifier from a list created from a database, to ensure consistency between the ID's in the database and those in the event tables.

the creation of an event database that very accurately displays the surface water entities being mapped, without altering the underlying nationally consistent coverage.

The Reach Indexing Tool (RIT), developed by EPA OWOW, takes advantage of the dynamic segmentation model and allows the user to input data in a user-friendly, menu-driven environment (Figure 7). The concept behind the RIT is that the user would extract a list of surface waterbody entities they wish to index (like 305(b) waterbody segments). The Reach Indexing Tool then allows the user to select the appropriate reach segments in RF3, add the segments to the event table, and choose an identifier from the list created from the surface water database. The RIT also allows the user to modify the start and end points of the event segments.

RESULTS

Over 18,000 303(d) listed waters were located and georeferenced for the 1998 303(d) reporting cycle. The data products created from the reach indexing initiative can be viewed and downloaded from http://www.epa.gov/owow/tmdl/index.html. Two national shapefiles (linear and point) created from the reach indexing event tables are available for download. Once the 303(d) waters were georeferenced to a hydrography coverage, it was possible to calculate size estimates for these waters. These sizes were used for the national summary in Figure 8. Once mapped, the 303(d) waters could also be linked to EPA's TMDL Tracking System and the 303(d) attribute data, so features like causes and sources of impairment could be mapped and analyzed (Figure 9).

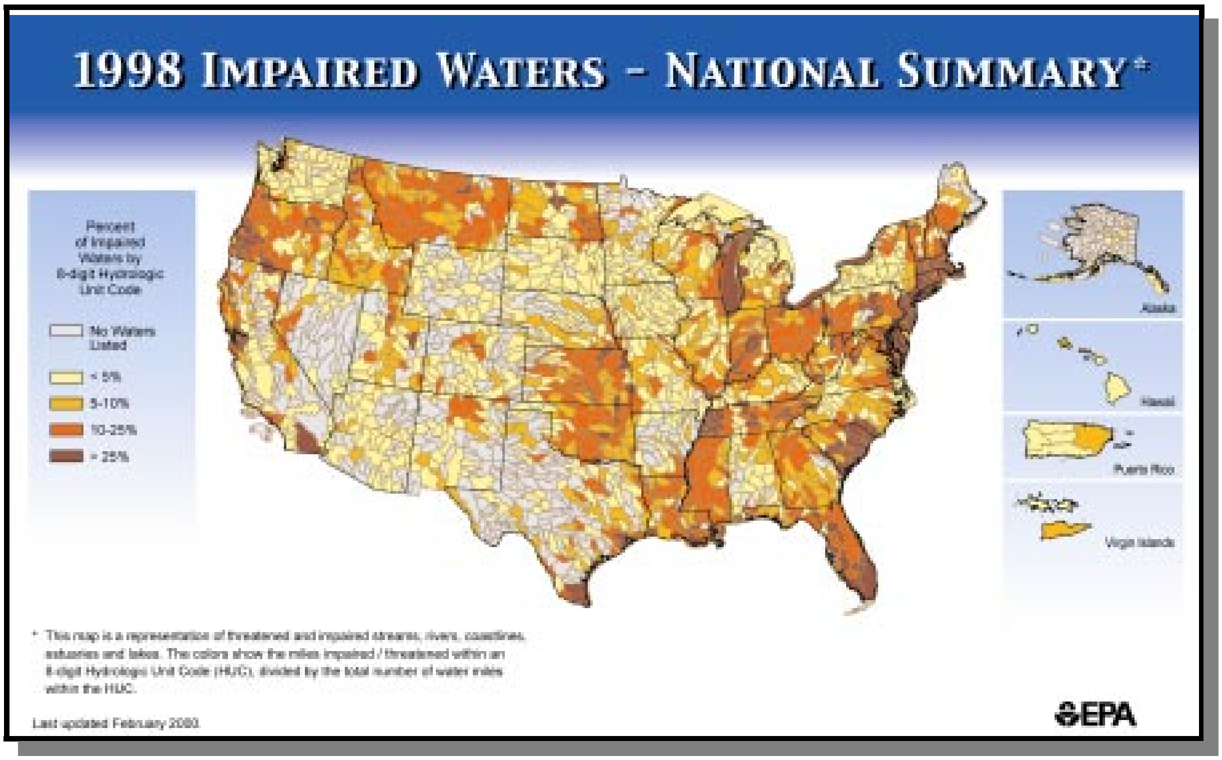

Figure 8. 1998 Percent of Impaired Waters by 8-digit USGS Cataloging Unit.

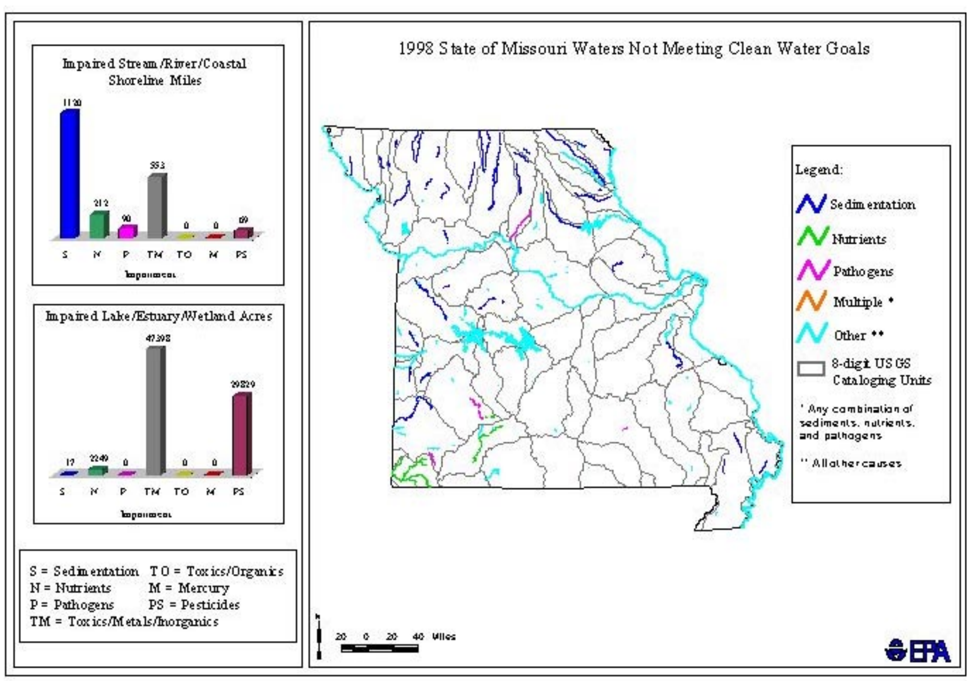

Figure 9. Information from the TMDL Tracking System displayed using the georeferencing work created by the RIT.

DISCUSSION

In addition to 303(d), reach indexing has been performed for several state 305(b) data sets. An initiative is also currently under way to reach index state water quality standards. Each program has an associated database that stores the attributes for the programmatic entities. These data sets and the event tables created through reach indexing will be brought together in a new EPA Oracle-based system called WATERs (for Watershed Assessment Tracking Environmental Results) (Figure 10) in FY2001. WATERs will be a web-based, interactive, GIS query tool that will allow users to query and display attribute information for all of the component surface water programs. The ability to query and map water quality information are great products of the indexing effort but the largest benefit will accrue when core water data sets are integrated. There is not always a direct relationship between the entities that states delineate for their surface water programs (like water quality standards, 305(b) and 303(d)), but WATERs will be able to capitalize on the link provided through georeferencing all of the entities to the National Hydrography Dataset. The system will be able to identify how these entities overlay and return the appropriate information within a hydrographic context. An example query of WATERs

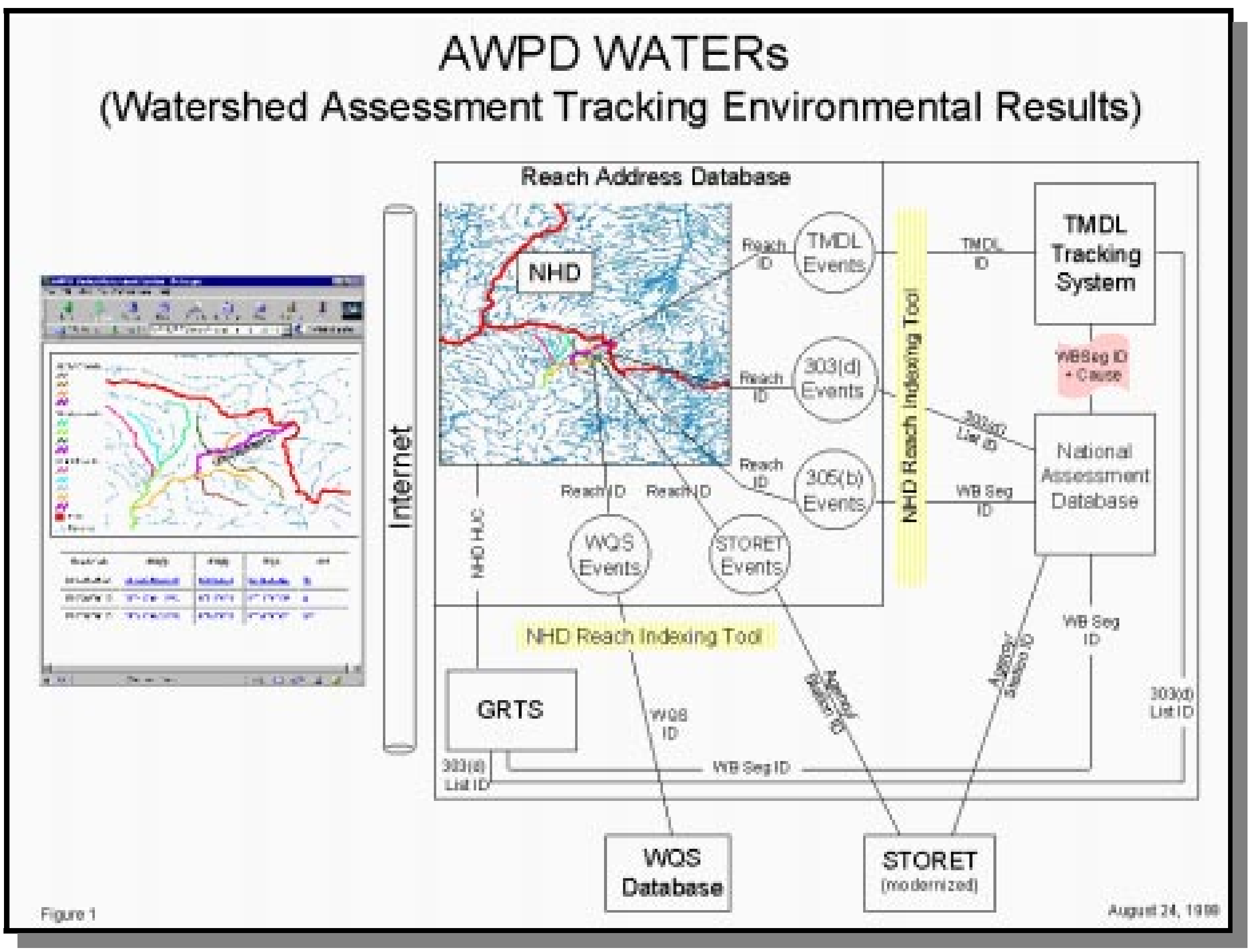

Figure 10. The relationship between the components of the WATERs system. All of the event tables created with the RIT, become part of the Reach Addressing Database, which serves as a common bond between the WATERs components (Diagram developed by the WATERs team including EPA, INDUS Corp., and RTI).

would be “show me all 303(d) listed waters for temperature on streams with a designated use of cold water fishery and within ten miles of any riparian best management practices.”

The WATERs system will be a large-scale application that will be useful to EPA Headquarters, Regional Offices, states, and the public. Users will be able to download subsets of surface water data and associated georeferencing data for a user-defined extent. These data can then be overlain and analyzed with other data sets of interest to the user.

The RIT application can be downloaded from http://georef.rit.org and is available to users who wish to manage their georeferencing data using a national hydrography coverage. The new National Hydrography Dataset-Reach Indexing Tool (NHD-RIT) not only allows users to delineate waterbodies in an associated database, but now users can take advantage of an additional field in the event table that allows them to store an attribute value. So, if users want only to map a certain attribute, like sediment type, they do not need to develop an outside database. By mapping attribute data like sediment type with the NHD-RIT, they can use these data directly with the water quality data downloaded from WATERs.

The ability to link water quality data to accurate, spatially determinate water quality coverages through WATERs provides a very powerful querying and mapping capability to interested users. Improvements in the way water quality data are managed and the improvements in the resolution of the data offered through NHD will help state and federal managers better identify areas of concern and enable them to allocate their resources appropriately.

Acknowledgments

The work described in this paper was funded by the U.S. Environmental Protection Agency under Contract 68-C7-0056 with Research Triangle Institute (RTI). RTI gratefully acknowledges this support.

Disclaimer: This paper has been reviewed in accordance with the U.S. Environmental Protection Agency’s peer and administrative review policies and approved for presentation and publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

Cooter, W., J. Fountain, P. Iliev, W. Wheaton, and R. Dodd. 1996. Making the most of state and tribal water quality assessment data. In *Watershed ’96: Proceedings of a National Conference on Watershed Management*, Baltimore, MD, June 8-12, 1996. Alexandria, VA: Water Environment Federation.

Cooter, W., J.M. McCarthy, and P. Iliev. 1998. State and tribal assessment databases: applications as watershed indicators and management tools. In: *Watershed '98: Proceedings of a National Conference on Watershed Management*, Denver, CO, May 3-6, 1998. Alexandria, VA: Water Environment Federation.

Dewald, T.G., C. McKay, S. Hansen, and W. Wheaton 1996. Managing watershed data with the USEPA Reach File. In: *Watershed '96: Proceedings of a National Conference on Watershed Management*, Baltimore, MD, June 8-12, 1996. Alexandria, VA: Water Environment Federation.